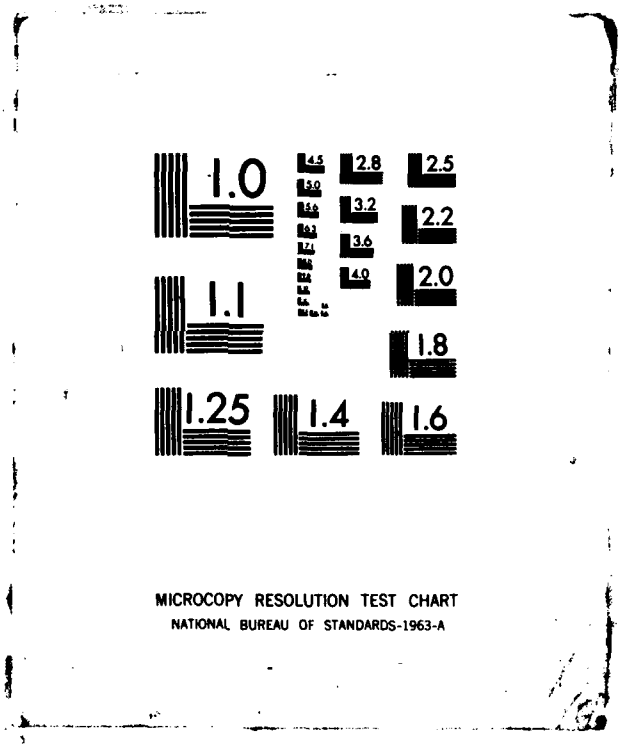
AD-A123 903 A CLASS OF GENERALIZED CORRELATION COEFFICIENTS(U) 1/1
FLORIDA STATE UNIV TALLAHASSEE DEPT OF STATISTICS
P LIN NOV 81 FSU-STATISTICS-M601 N00014-80-C-0093
UNCLASSIFIED F/G 12/1 NL

END
DATE FILMED
83
DTIC


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0
4.5
5.0
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

A Class of Generalized Correlation Coefficients[1]

by

Pi-Erh Lin

A Class of Generalized Correlation Coefficients[1]

by

Pi-Erh Lin

FSU Statistics Report No. M601
ONR Technical Report No. 156

November, 1981

The Florida State University
Department of Statistics
Tallahassee, Florida 32306

[1]This work was supported by the Army, Navy and Air Force under Office of Naval Research Contract No. N00014-80-C-0093. Reproduction in whole or in part is permitted for any purpose of the United States Government.

A Class of Generalized Correlation Coefficients

by

Pi-Erh Lin

Department of Statistics, The Florida State University

SUMMARY

Let $\underset{\sim}{X} = (\underset{\sim}{X}^{(1)'}\underset{\sim}{X}^{(2)'})'$ be a $(p + q)$-variate random vector $(p \geq 1, q \geq 1)$ with $E\underset{\sim}{X} = \underset{\sim}{0}$ and $E\underset{\sim}{X}\underset{\sim}{X}' = \Sigma > 0$. Partition Σ into p- and q- rows and columns such that $E(\underset{\sim}{X}^{(i)}\underset{\sim}{X}^{(j)'}) = \Sigma_{ij}(i, j = 1, 2)$. Let $\lambda_1^2, \ldots, \lambda_s^2$ be the nonzero characteristic roots of $\Sigma_{11}^{-1}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$ where s is the rank of Σ_{12}. Based on these roots, a class of generalized correlation coefficients between $\underset{\sim}{X}^{(1)}$ and $\underset{\sim}{X}^{(2)}$ are obtained. Some optimal properties of a generalized correlation coefficient in the class are presented. Among various special cases, two information number related coefficients are derived. However, an attractive generalized correlation coefficient is given by $\rho_W = 1 - \prod_{i=1}^{p}(1 - \lambda_i^2)$ which is derived from the vector coefficient of alienation. A justification for the use of ρ_W is also included.

[1]The work was supported by the Army, Navy and Air Force under Office of Naval Research Contract No. N00014-80-C-0093. Reproduction in whole or in part is permitted for any purpose of the United States Government.

AMS 1970 subject classifications: Primary 62H20

Key Words and Phrases: Canonical correlation; Coefficient of alienation; Information number, Invariant with respect to a linear transformation; Measure of association.

1. INTRODUCTION

Let $\underset{\sim}{X}$ be a $(p + q)$-variate random vector $(p \geq 1, q \geq 1)$ with mean vector $\underset{\sim}{\mu}$ and covariance matrix Σ. Partition

$$\underset{\sim}{X} = \begin{pmatrix} \underset{\sim}{X}^{(1)} \\ \underset{\sim}{X}^{(2)} \end{pmatrix} \quad \text{and} \quad \underset{\sim}{\mu} = \begin{pmatrix} \underset{\sim}{\mu}^{(1)} \\ \underset{\sim}{\mu}^{(2)} \end{pmatrix}$$

into p- and q- component subvectors respectively. Assume, without loss of generality, that $p \leq q$. Define

$$\Sigma_{ij} = E(\underset{\sim}{X}^{(i)} - \underset{\sim}{\mu}^{(i)})(\underset{\sim}{X}^{(j)} - \underset{\sim}{\mu}^{(j)})', \quad i = 1, 2.$$

Then Σ can be partitioned into p- and q- rows and columns as

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}. \tag{1.1}$$

In this paper we will give a class of generalized correlation coefficients as a measure of association between $\underset{\sim}{X}^{(1)}$ and $\underset{\sim}{X}^{(2)}$. In particular, we will propose a generalized correlation with desirable optimal properties. For the convenience of presentation, it will be assumed that Σ is positive definite throughout the study. However, if Σ is positive semidefinite the results will remain true with Σ_{ii}^{-1} replaced by the Moore-Penrose generalized inverse Σ_{ii}^{+}, $i = 1, 2$.

For the case of two random variables X_1 and X_2 (say), i.e., $p = q = 1$, most frequently used correlations between X_1 and X_2, such as the Pearson product-moment correlation and the Spearman rank correlation, are invariant with respect to location and scale changes. It

seems reasonable then to restrict our attention to those measures of association between $\underset{\sim}{X}^{(1)}$ and $\underset{\sim}{X}^{(2)}$ which are invariant with respect to a location change and a nonsingular linear transformation. More specifically, let $\underset{\sim}{d}^{(1)}$ be a $p \times 1$ vector and $\underset{\sim}{d}^{(2)}$ be a $q \times 1$ vector, and let B be a $p \times p$ nonsingular matrix and C be a $q \times q$ nonsingular matrix. Then consider the transformation on $\underset{\sim}{X}^{(1)}$ and $\underset{\sim}{X}^{(2)}$ given by

$$\underset{\sim}{X}^{(1)} \to B\underset{\sim}{X}^{(1)} + \underset{\sim}{d}^{(1)}$$
$$\underset{\sim}{X}^{(2)} \to C\underset{\sim}{X}^{(2)} + \underset{\sim}{d}^{(2)}.$$

The above transformation induces a transformation on the parameter space $(\underset{\sim}{\mu}, \Sigma)$ given by

$$\begin{aligned} \underset{\sim}{\mu}^{(1)} &\to B\underset{\sim}{\mu}^{(1)} + \underset{\sim}{d}^{(1)} \\ \underset{\sim}{\mu}^{(2)} &\to C\underset{\sim}{\mu}^{(2)} + \underset{\sim}{d}^{(2)} \\ \Sigma_{11} &\to B\Sigma_{11}B' \\ \Sigma_{12} &\to B\Sigma_{12}C' \\ \Sigma_{22} &\to C\Sigma_{22}C' \end{aligned} \tag{1.2}$$

Thus, in the following section, we will characterize a class of measures of association between $\underset{\sim}{X}^{(1)}$ and $\underset{\sim}{X}^{(2)}$ which are invariant with respect to the transformation given by (1.2). Desirable properties for these measures will be studied. In Section 3, two measures of association will be obtained from the Kullback and Matsusita information numbers. Finally, in Section 4 a special measure, a generalized correlation, will be proposed and its justification made.

2. A CLASS OF MEASURES OF ASSOCIATION

A measure of association between $\underset{\sim}{X}^{(1)}$ and $\underset{\sim}{X}^{(2)}$ is a real number between 0 and 1. It measures a certain relationship between $\underset{\sim}{X}^{(1)}$ and $\underset{\sim}{X}^{(2)}$ such that the larger the measure is the stronger the relationship will be, and vice versa. Through the following lemma we will see that the nonzero characteristic roots of $\Sigma_{11}^{-1}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$, denoted by $\lambda_i^2 = \mathrm{ch}_i(\Sigma_{11}^{-1}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21})$, $i = 1, \ldots, s$, play an important role in defining such a measure of association.

LEMMA 1. Let $\lambda_1^2 \geq \ldots \geq \lambda_s^2 > 0$ be the nonzero characteristic roots of $\Sigma_{11}^{-1}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{22}$ where s is the rank of Σ_{12}. Then any function of $(\underset{\sim}{\mu}, \Sigma)$ which is invariant with respect to the transformation given by (1.2) is a function of $(\lambda_1^2, \ldots, \lambda_s^2)$.

PROOF. Let f be such a function. Choose nonsingular matrices B and C in (1.2) such that $B\Sigma_{11}B' = I_p$, $C\Sigma_{22}C' = I_q$ and let $\underset{\sim}{d}^{(1)} = -B\underset{\sim}{\mu}^{(1)}$ and $\underset{\sim}{d}^{(2)} = -C\underset{\sim}{\mu}^{(2)}$. Then

$$f(\underset{\sim}{\mu}^{(1)}, \underset{\sim}{\mu}^{(2)}, \Sigma_{11}, \Sigma_{22}, \Sigma_{12})$$

$$= f(\underset{\sim}{0}, \underset{\sim}{0}, I_p, I_q, B\Sigma_{12}C')$$

$$= f_1(B\Sigma_{12}C').$$

But since $B\Sigma_{12}C'$ is a $p \times q$ matrix of rank s, there exists an orthogonal matrix Q of order q such that

$$B\Sigma_{12}C'Q' = (T \vdots 0)$$

where T is a $p \times p$ lower triangular matrix of rank s with nonnegative diagonal elements and 0 is a $p \times (q - p)$ matrix of zero. It is clear that T is a function of $TT' = B\Sigma_{12}C'C\Sigma_{21}B'$ and that there exists an orthogonal matrix P of order p such that

$$PTT'P' = \Delta = \mathrm{diag}(\delta_1, \ldots, \delta_s, 0, \ldots, 0)$$

with $\delta_1 \geq \ldots \geq \delta_s > 0$. However, for $i = 1, \ldots, s$,

$$\begin{aligned}\delta_i &= \mathrm{ch}_i(B\Sigma_{12}C'C\Sigma_{21}B') \\ &= \mathrm{ch}_i(B'B\Sigma_{12}C'C\Sigma_{21}) \\ &= \mathrm{ch}_i(\Sigma_{11}^{-1}\Sigma_{12}\ \Sigma_{22}^{-1}\Sigma_{21}) = \lambda_i^2.\end{aligned}$$

Therefore,

$$\begin{aligned}f_1(B\Sigma_{12}C') &= f_2(B\Sigma_{12}C'C\Sigma_{21}B') \\ &= f_2(\Delta) \\ &= f_3(\delta_1, \ldots, \delta_s) \\ &= f_3(\lambda_1^2, \ldots, \lambda_s^2),\end{aligned}$$

completing the proof. □

In fact, various authors have proposed measures of association based on $\lambda_1^2, \ldots, \lambda_s^2$. For example, Zhang (1978) studies the following five measures of association

i) $\rho^{(1)} = \sum_{i=1}^{s} \lambda_i^2/s$ (arithmetic mean)

(ii) $\rho^{(2)} = (\prod_{i=1}^{s} \lambda_i^2)^{1/s}$ (geometric mean)

(iii) $\rho^{(3)} = (\sum_{i=1}^{s} \lambda_i^{-2}/s)^{-1}$ (harmonic mean)

(iv) $\rho^{(4)} = \lambda_1^2$ $(\max_{1 \le i \le s} \{\lambda_i^2\})$

and

(v) $\rho^{(5)} = \lambda_s^2$. $(\min_{1 \le i \le s} \{\lambda_i^2\})$

It is clear that $0 \le \rho^{(5)} \le \rho^{(3)} \le \rho^{(2)} \le \rho^{(1)} \le \rho^{(4)} \le 1$. Recently, Jupp and Mardia (1980), in a study of correlation for directional data, propose to use $\mathrm{tr}(\Sigma_{11}^{-1}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21})$ as a measure of association which, of course, is equivalent to $\rho^{(1)}$. Now we are in a position to propose a general class of measures of association between $\underset{\sim}{X}^{(1)}$ and $\underset{\sim}{X}^{(2)}$, based on $\lambda_1^2, \ldots, \lambda_s^2$.

DEFINITION. Let h be a strictly increasing function mapping [0, 1] onto itself such that $h(0) = 0$ and $h(1) = 1$, and let g be a strictly monotone function mapping [0, 1] onto [a, b], $0 \le a < b \le \infty$, such that either

i) $g(0) = a$ and $g(1) = b$, or

ii) $g(0) = b$ and $g(1) = a$.

Then a generalized correlation coefficient, $R_{12}(g, h)$, between $\underset{\sim}{X}^{(1)}$ and $\underset{\sim}{X}^{(2)}$ is defined by

$$R_{12}(g, h) = g^{-1}\{\sum_{i=1}^{s} c_i g[h(\lambda_i^2)]\} \text{ if } s > 0$$

$$= 0 \text{ if } s = 0$$

where $\lambda_1^2, \ldots, \lambda_s^2$ are the nonzero characteristic roots of $\Sigma_{11}^{-1}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$ and $c_i \geq 0$ with $\sum_{i=1}^{s} c_i = 1$. If $a = 0$ and $b = \infty$, then $\sum_{i=1}^{s} c_i = 1$ is not required.

Note that the set C of all such generalized correlation coefficients, $R_{12}(g, h)$, characterizes a class of measures of association between $\underset{\sim}{X}^{(1)}$ and $\underset{\sim}{X}^{(2)}$. A generalized correlation in C, with given g and h functions, possesses the following desirable properties:

(1) $R_{12}(g, h) = R_{21}(g, h)$. This follows from the fact that the ith largest characteristic root of $\Sigma_{11}^{-1}\Sigma_{12}\,\Sigma_{22}^{-1}\,\Sigma_{21}$ is the same as the ith largest characteristic root of $\Sigma_{22}^{-1}\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}$, $i = 1, \ldots, s$. Thus a generalized correlation between $\underset{\sim}{X}^{(1)}$ and $\underset{\sim}{X}^{(2)}$ is the same as that between $\underset{\sim}{X}^{(2)}$ and $\underset{\sim}{X}^{(1)}$.

(2) $0 \leq R_{12}(g, h) \leq 1$. The characteristic roots of $\Sigma_{11}^{-1}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$ are bounded between 0 and 1. This implies that, for $i = 1, \ldots, s$, $h(\lambda_i^2) \in [0, 1]$, $g[h(\lambda_i^2)] \in [a, b]$, and that $\sum_{i=1}^{s} c_i g[h(\lambda_i^2)] \in [a, b]$. Thus $g^{-1}\{\sum_{i=1}^{s} c_i g[h(\lambda_i^2)]\} \in [0, 1]$.

(3). $R_{12}(g, h) = 0$ if and only if $\Sigma_{12} = 0$. If $\Sigma_{12} = 0$, then $s = 0$ and, by definition, $R_{12}(g, h) = 0$. Conversely, if $R_{12}(g, h) = 0$, then $s = 0$ and $0 = (\Sigma_{11}^{-\frac{1}{2}}\Sigma_{12}\Sigma_{22}^{-\frac{1}{2}})(\Sigma_{11}^{-\frac{1}{2}}\Sigma_{12}\Sigma_{22}^{-\frac{1}{2}})'$. This implies that $\Sigma_{11}^{-\frac{1}{2}}\Sigma_{12}\Sigma_{22}^{-\frac{1}{2}} = 0$ and hence $\Sigma_{12} = 0$.

(4). If there exists a $p \times q$ matrix H of rank p such that $\underset{\sim}{X}^{(1)} = H\underset{\sim}{X}^{(2)}$, then $R_{12}(g, h) = 1$. If $\underset{\sim}{X}^{(1)} = H\underset{\sim}{X}^{(2)}$ then $\Sigma_{11} = H\Sigma_{22}H'$, $\Sigma_{12} = H\Sigma_{22}$, and

$$\Sigma_{11}^{-1}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} = (H\Sigma_{22}H')^{-1}H\Sigma_{22}\Sigma_{22}^{-1}\Sigma_{22}H' = (H\Sigma_{22}H')^{-1}H\Sigma_{22}H' = I_p.$$

Thus $\lambda_1^2 = \ldots = \lambda_p^2 = 1$. Therefore $R_{12}(g, h) = g^{-1}[g(1)\sum_{i=1}^{p} c_i] = 1$.

(5) $R_{12}(g, h)$ is monotone nondecreasing in each λ_i^2, $i = 1, \ldots, s$.

Since

$$g[R_{12}(g, h)] = \sum_{i=1}^{s} c_i h[h(\lambda_i^2)] ,$$

it follows that

$$\partial R_{12}(g, h)/\partial\lambda_i^2 = c_i g'[h(\lambda_i^2)]h'(\lambda_i^2)/g'[R_{12}(g, h)] \geq 0$$

for $i = 1, \ldots, s$.

(6). $R_{12}(g, h)$ between $\underset{\sim}{X}^{(1)}$ and $\underset{\sim}{X}^{(2)}$ is the same as that between $(\underset{\sim}{X}^{(1)'}\underset{\sim}{Y}^{(1)'})'$ and $(\underset{\sim}{X}^{(2)'}\underset{\sim}{Y}^{(2)'})'$ for any $m \times 1$ and $n \times 1$ uncorrelated random vectors $\underset{\sim}{Y}^{(1)}$ and $\underset{\sim}{Y}^{(2)}$ which are uncorrelated with $\underset{\sim}{X}^{(1)}$ and $\underset{\sim}{X}^{(2)}$.

Let Φ_{ii} be the covariance matrix of $\underset{\sim}{Y}^{(i)}$, $i = 1, 2$. Define $\underset{\sim}{W}^{(i)} = (\underset{\sim}{X}^{(i)'}\underset{\sim}{Y}^{(i)'})'$, $i = 1, 2$, and let Σ^* be the covariance matrix of $(\underset{\sim}{W}^{(1)'}\underset{\sim}{W}^{(2)'})'$. Partition Σ^* into $(p + m)$- and $(q + n)$- rows and columns as

$$\Sigma^* = \begin{pmatrix} \Sigma_{11}^* & \Sigma_{12}^* \\ \Sigma_{21}^* & \Sigma_{22}^* \end{pmatrix}$$

where

$$\Sigma_{ii}^* = \begin{pmatrix} \Sigma_{ii} & 0 \\ 0 & \Phi_{ii} \end{pmatrix} , \quad i = 1, 2,$$

and

$$\Sigma_{12}^* = \begin{pmatrix} \Sigma_{12} & 0 \\ 0 & 0 \end{pmatrix} = \Sigma_{21}^{*'}$$

Assume, without loss of generality, that Φ_{11} and Φ_{22} are positive definite matrices. Then

$$\Sigma_{11}^{*-1}\Sigma_{12}^{*}\Sigma_{22}^{*-1}\Sigma_{21}^{*} = \begin{pmatrix} \Sigma_{11}^{-1}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} & 0 \\ 0 & 0 \end{pmatrix}$$

and thus the nonzero characteristic roots of $\Sigma_{11}^{*-1}\Sigma_{12}^{*}\Sigma_{22}^{*-1}\Sigma_{21}^{*}$ are the same as those of $\Sigma_{11}^{-1}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$. Hence $R_{12}(g, h)$ between $\underset{\sim}{X}^{(1)}$ and $\underset{\sim}{X}^{(2)}$ remains unchanged when the subvector $\underset{\sim}{X}^{(i)}$ is augmented by an uncorrelated vector $\underset{\sim}{Y}^{(i)}$, $i = 1, 2$, where $\underset{\sim}{Y}^{(1)}$ and $\underset{\sim}{Y}^{(2)}$ are uncorrelated between themselves.

Note that the five measures of association studied by Zhang (1978) are special cases of $R_{12}(g, h)$. This can be verified by properly identifying the g and h functions. For example, (i) $\rho^{(1)}$ is obtained by letting $h(x) = x$, $g(y) = y$, and $c_i = 1/s$; (ii) $\rho^{(2)}$ is obtained by letting $h(x) = x$, $g(y) = -\ln y$, and $c_i = 1/s$; and so on. For all of the five measures, the h function is always the identity function, i.e., $h(x) = x$. In the following section, we will present two measures in C where the h function is not the identity function.

3. MEASURES BASED ON INFORMATION NUMBERS

In the previous section, a large class C of generalized correlations between $\underset{\sim}{X}^{(1)}$ and $\underset{\sim}{X}^{(2)}$ is defined without specifying the joint distribution of $\underset{\sim}{X}^{(1)}$ and $\underset{\sim}{X}^{(2)}$ so long as $\Sigma_{ij}(i, j = 1, 2)$ exist. In the remainder of the study, we will assume that $\underset{\sim}{X}$ has a $(p + q)$-variate normal distribution with mean vector $\underset{\sim}{\mu}$ and covariance matrix Σ. Assume, without loss of generality, that $\underset{\sim}{\mu} = \underset{\sim}{0}$ and partition Σ as in (1.1). Then the multivariate multiple regression of $\underset{\sim}{X}^{(1)}$ on $\underset{\sim}{X}^{(2)}$ is given by

$$\hat{\underset{\sim}{X}}^{(1)} = \Sigma_{12}\Sigma_{22}^{-1}\underset{\sim}{X}^{(2)} \tag{3.1}$$

Usually, $\hat{\underset{\sim}{X}}^{(1)}$ is used to predict the value of $\underset{\sim}{X}^{(1)}$ and the precision of such a prediction may be evaluated by the use of various information numbers. In this section we will obtain two measures of association between $\underset{\sim}{X}^{(1)}$ and $\underset{\sim}{X}^{(2)}$, which are members of the general class C, based on the Kullback and Matsusita information numbers between the distributions of $\underset{\sim}{X}^{(1)}$ and $\hat{\underset{\sim}{X}}^{(1)}$.

3.1 KULLBACK INFORMATION NUMBER

For ease of presentation, let $\underset{\sim}{Z}^{(2)} = \hat{\underset{\sim}{X}}^{(1)}$ and let $f_1(\underset{\sim}{x})$ and $f_2(\underset{\sim}{x})$ denote the density functions of $\underset{\sim}{X}^{(1)}$ and $\underset{\sim}{Z}^{(2)}$ respectively assuming Σ_{12} is of full rank. Then the Kullback information number between $\underset{\sim}{X}^{(1)}$ and $\underset{\sim}{Z}^{(2)}$ is given by

$$K(2, 1) = E_2\{ln[f_2(\underset{\sim}{X})/f_1(\underset{\sim}{X})]\}$$

where E_2 denotes the expectation taken with respect to the distribution of $\underset{\sim}{Z}^{(2)}$. It follows that

$$K(2, 1) = -(1/2)\ln|\Sigma_{11}^{-1}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}| - p/2 + (1/2)\mathrm{tr}(\Sigma_{11}^{-1}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21})$$
$$= -(1/2)\sum_{i=1}^{p}(\ln\lambda_i^2 + 1 - \lambda_i^2).$$

Thus a measure of association based on K(2, 1) may be defined as

$$\rho_{K(2,1)} \equiv \exp[-K(2, 1)]$$
$$= \exp[(1/2)\sum_{i=1}^{p}(\ln\lambda_i^2 + 1 - \lambda_i^2)]$$
$$= \prod_{i=1}^{p}\exp[(1/2)(\ln\lambda_i^2 + 1 - \lambda_i^2)].$$

This is a member of $\mathcal{C}$ as can be verified by taking

$$h(x) = \exp(\ln x + 1 - x)$$
$$g(y) = -\ln y$$
and $$c_i = 1/2,\ i = 1, \ldots, p.$$

3.2. MATSUSITA INFORMATION NUMBER

The Matsusita information number between $\underset{\sim}{X}^{(1)}$ and $\underset{\sim}{Z}^{(2)}$ is defined as

$$M(2, 1) \equiv \int_{R^p}[f_1^{\frac{1}{2}}(\underset{\sim}{x}) - f_2^{\frac{1}{2}}(\underset{\sim}{x})]^2 d\underset{\sim}{x}$$
$$= 2\{1 - \int_{R^p}[f_1(\underset{\sim}{x})f_2(\underset{\sim}{x})]^{\frac{1}{2}}d\underset{\sim}{x}\}$$

where R^p denotes the p-dimensional Euclidean space. It can be shown that

$$\int_{R^p}[f_1(\underset{\sim}{x})f_2(\underset{\sim}{x})]^{\frac{1}{2}}d\underset{\sim}{x} = \prod_{i=1}^{p}\left[\frac{2\lambda_i}{1 + \lambda_i^2}\right]^{\frac{1}{2}}.$$

Thus a measure of association between $\underset{\sim}{X}^{(1)}$ and $\underset{\sim}{X}^{(2)}$ based on M(2, 1)

may be defined as

$$\rho_{M(2,1)} = 1 - (1/2)M(2, 1)$$

$$= \prod_{i=1}^{p} \left(\frac{2\lambda_i}{1 + \lambda_i^2} \right)^{\frac{1}{2}}.$$

Taking

$$h(x) = \frac{2x^{\frac{1}{2}}}{1 + x}$$

$$g(y) = -\ln y$$

and

$$c_i = 1/2, \; i = 1, \ldots, p,$$

it is easily seen that $\rho_{M(2,1)}$ is a generalized correlation in the class C.

More measures of association between $\underset{\sim}{X}^{(1)}$ and $\underset{\sim}{X}^{(2)}$ may be obtained using other information numbers in the same manner as those presented above.

4. MEASURE BASED ON COEFFICIENT OF ALIENATION

Finally, in this section, we will propose a special generalized correlation in C through the use of the vector coefficient of alienation. This is a direct generalization of the square of the Pearson correlation between two normal random variables. More specifically, let $(X_1, X_2)'$ have a bivariate normal distribution with mean vector $(\mu_1, \mu_2)'$ and covariance matrix $\Sigma = (\sigma_{ij})$, $i, j = 1, 2$. Then the coefficient of alienation is given by

$$\frac{\mathrm{Var}(X_1|x_2)}{\mathrm{Var}(X_1)} = \frac{\sigma_{11} - \sigma_{12}^2/\sigma_{22}}{\sigma_{11}} = 1 - \rho^2$$

where ρ is the Pearson correlation between X_1 and X_2. In the case of two subvectors $\underset{\sim}{X}^{(1)}$ and $\underset{\sim}{X}^{(2)}$, the vector coefficient of alienation is given by [see, e.g., Anderson (1958), p. 244]

$$\frac{|\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}|}{|\Sigma_{11}|} = \prod_{i=1}^{s} (1 - \lambda_i^2) \tag{4.1}$$

where $\lambda_1^2, \ldots, \lambda_s^2$ are the nonzero characteristic roots of $\Sigma_{11}^{-1}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$. Thus a measure of association between $\underset{\sim}{X}^{(1)}$ and $\underset{\sim}{X}^{(2)}$, based on (4.1), may be defined as

$$\rho_W = 1 - \prod_{i=1}^{s} (1 - \lambda_i^2),$$

or, equivalently, as

$$\rho_W = 1 - \prod_{i=1}^{p} (1 - \lambda_i^2) \tag{4.2}$$

since $\lambda_{s+1}^2 = \ldots = \lambda_p^2 = 0$.

Taking $h(x) = x$, $g(y) = -\ln(1 - y)$, and $c_i = 1$, $i = 1, \ldots, p$, it follows that ρ_W is a generalized correlation in the class C and hence enjoys all of the properties discussed in Section 2. Furthermore, because of the special choice of the g and h functions, property (4) may now be strengthened for ρ_W as follows:

(4') $\rho_W = 1$ if and only if there exists $\underset{\sim}{\alpha} \in R^p$ and $\underset{\sim}{\beta} \in R^q$ such that $\underset{\sim}{\alpha}'\underset{\sim}{X}^{(1)} = \underset{\sim}{\beta}'\underset{\sim}{X}^{(2)}$ and $s \geq 1$. The fact that $\underset{\sim}{\alpha}'\underset{\sim}{X}^{(1)} = \underset{\sim}{\beta}'\underset{\sim}{X}^{(2)}$ is equivalent to the first canonical correlation λ_1 being 1; this, in turn, is equivalent to $\rho_W = 1$.

We now turn to the statistical inference on the relationship between $\underset{\sim}{X}^{(1)}$ and $\underset{\sim}{X}^{(2)}$. This will provide a further justification for proposing the use of ρ_W as a reasonable measure of association between $\underset{\sim}{X}^{(1)}$ and $\underset{\sim}{X}^{(2)}$.

Let $\underset{\sim}{X}_1, \ldots, \underset{\sim}{X}_N$ be a sample of size N from $N_{p+q}(\underset{\sim}{\mu}, \sum)$ and let

$$A = \sum_{k=1}^{N} (\underset{\sim}{X}_k - \overline{\underset{\sim}{X}})(\underset{\sim}{X}_k - \overline{\underset{\sim}{X}})',$$

where $\overline{\underset{\sim}{X}} = (1/N)\sum_{k=1}^{N} \underset{\sim}{X}_k$. Partition A into p- and q- rows and columns as

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}.$$

It is well known that the maximum likelihood estimator of $\sum_{ij}$ is given

by $(1/N)A_{ij}$, $i = 1, 2$, and hence a natural estimator for $R_{12}(g, h)$ is given by

$$\hat{R}_{12}(g, h) = g^{-1}\{\sum_{i=1}^{\hat{s}} \hat{c}_i g[h(\hat{\lambda}_i^2)]\} \tag{4.3}$$

where $\hat{\lambda}_1^2, \ldots, \hat{\lambda}_{\hat{s}}^2$ are the nonzero characteristic roots of $A_{11}^{-1}A_{12}A_{22}^{-1}A_{21}$, $\hat{s}$ is the rank of A_{12}, and $\hat{c}_i$ is a function of $\hat{s}$. Since $(1/N)A_{ij} \to \Sigma_{ij}$ $(i, j = 1, 2)$ in probability as $N \to \infty$, it follows that $\hat{\lambda}_i^2 \to \lambda_i^2$, $\hat{s} \to s$, $\hat{c}_i \to c_i$ and that $\hat{R}_{12}(g, h) \to R_{12}(g, h)$, in probability, as $N \to \infty$. In particular, a natural estimator for ρ_W is given by

$$\hat{\rho}_W = 1 - \prod_{i=1}^{p} (1 - \hat{\lambda}_i^2), \tag{4.4}$$

where $\hat{\lambda}_1^2, \ldots, \hat{\lambda}_p^2$ are the characteristic roots of $A_{11}^{-1}A_{12}A_{22}^{-1}A_{21}$. Moreover, the estimator given by (4.4) is a function of the Wilks likelihood ratio statistic $|A|/(|A_{11}|\ |A_{22}|)$ for testing the hypothesis of independence between $\underset{\sim}{X}^{(1)}$ and $\underset{\sim}{X}^{(2)}$, i.e.,

$$\hat{\rho}_W = 1 - \frac{|A|}{|A_{11}|\ |A_{22}|} . \tag{4.5}$$

When N is sufficiently large, the null distribution of $-[N - 1 - (p + q + 1)/2]\ln(1 - \hat{\rho}_W)$ may be approximated by a chi-square distribution with pq degrees of freedom which is customarily used as a step-down procedure for the test of significance on the canonical correlations between $\underset{\sim}{X}^{(1)}$ and $\underset{\sim}{X}^{(2)}$. After the effects of the first, second, ..., mth $(m \leq p - 1)$ canonical-variate pairs have cumulatively been removed, we may define a "remaining" generalized correlation coefficient

between $\underset{\sim}{X}^{(1)}$ and $\underset{\sim}{X}^{(2)}$ as

$$\rho_{W(m+1,\ldots,p)} = 1 - \prod_{i=m+1}^{p} (1 - \lambda_i^2) \tag{4.6}$$

and its natural estimator as

$$\hat{\rho}_{W(m+1,\ldots,p)} = 1 - \prod_{i=m+1}^{p} (1 - \hat{\lambda}_i^2) . \tag{4.7}$$

Since $-[N - 1 - (p + q + 1)/2] \ln[1 - \hat{\rho}_{W(m+1,\ldots,p)}]$ is asymptotically distributed as χ^2 with $(p - m)(q - m)$ degrees of freedom when $\Sigma_{12} = 0$, it may be used to test whether a significant relationship still exists between $\underset{\sim}{X}^{(1)}$ and $\underset{\sim}{X}^{(2)}$ after the effects of the first m canonical-variate pairs have been removed. In the canonical correlation analysis, there are other test procedures available. For example,

(i) Lawley-Hotelling's trace: $\sum_{i=1}^{p} [\lambda_i^2/(1 - \lambda_i^2)]$ (4.8)

and

(ii) Pillai's trace: $\sum_{i=1}^{s} \lambda_i^2$. (4.9)

Based on (4.8) and (4.9), generalized correlation coefficients, which are members of the class C, may be obtained. However, they are not as intuitively appealing as ρ_W , which is discussed above.

REFERENCES

ANDERSON, T.W. (1958). An Introduction to Multivariate Statistical Analysis. New York: Wiley.

JUPP, P.E. and MARDIA, K. V. (1980). A general correlation coefficient for directional data and related regression problems. Biometrika 67 163-173.

ZHANG, Y.T. (1978). Generalized correlation coefficients and their applications. Acta Math. Appl. Sinica 1 312-320.

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1. REPORT NUMBER FSU Report No M601 ONR Report No 156	2. GOVT ACCESSION NO. AD-A123903
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and subtitle) A Class of Generalized Correlation Coefficients	5. TYPE OF REPORT & PERIOD COVERED Technical
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Pi-Erh Lin	8. CONTRACT OR GRANT NUMBER(s) ONR Contract No N00014-80-C-0093
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Statistics Florida State University Tallahassee, FL. 32306	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Statistics and Probability Program Arlington, VA. 22217	12. REPORT DATE November, 1981 13. NUMBER OF PAGES 16
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS. (of this report) 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this report)
Distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from report)

18. SUPPLEMENTARY NOTES

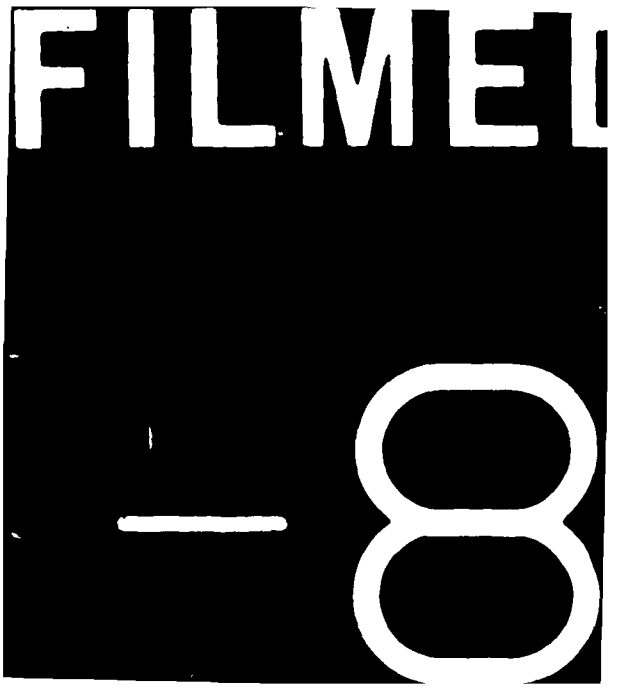
19. KEY WORDS
Canonical correlation; Coefficient of alienation; Information number, Invariant with respect to a linear transformation; Measure of association.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Let $\underset{\sim}{X} = (\underset{\sim}{X}^{(1)\prime}\ \underset{\sim}{X}^{(2)\prime})'$ be a $(p + q)$-variate random vector ($p \geq 1$, $q \geq 1$) with $E\underset{\sim}{X} = \underset{\sim}{0}$ and $E\underset{\sim}{X}\underset{\sim}{X}' = \Sigma > 0$. Partition Σ into p- and q- rows and columns such that $E(\underset{\sim}{X}^{(i)}\underset{\sim}{X}^{(j)\prime}) = \Sigma_{ij}(i, j = 1, 2)$. Let $\lambda_1^2, \ldots, \lambda_s^2$ be the nonzero characteristic roots of $\Sigma_{11}^{-1}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$ where s is the rank of Σ_{12}. Based on these roots, a class of generalized correlation coefficients between $\underset{\sim}{X}^{(1)}$ and $\underset{\sim}{X}^{(2)}$ are obtained. Some optimal properties of a generalized correlation coefficient in the class are presented. Among various special cases, two information number related

20. ABSTRACT continuation

coefficients are derived. However, an attractive generalized correlation coefficient is given by $\rho_W = 1 - \prod_{i=1}^{p} (1 - \lambda_i^2)$ which is derived from the vector coefficient of alienation. A justification for the use of ρ_W is also included.


FILME
8